THIRD EDITION

HANDBOOK OF RESEARCH ON TEACHING THE ENGLISH LANGUAGE ARTS

SPONSORED BY THE INTERNATIONAL READING ASSOCIATION & THE NATIONAL COUNCIL OF TEACHERS OF ENGLISH

EDITED BY DIANE LAPP • DOUGLAS FISHER
The study of language and literacy over time has a long history. The history of child development studies involves three distinct sets that are based on time. Each examines reading and writing and those activities more broadly conceived as literacy and language arts over different units of time. The sets of studies range from a long-term time perspective over multiple years, through a medium-term frame of weeks, months, and years, to a short-term or micro genetic frame of seconds, minutes and hours. The latter two in part owe their features to the older set of studies with the long time frame. So the longer term studies provide an historically and conceptually appropriate introduction for the other two sets.

Classic Child Developmental Designs

In 1972 Clay presented the results of an ongoing study into children’s literacy development from the entrance to school. Initially over 1 year, the follow up of 100 children extended to a further 2 and 3 years. The design she used was a classic longitudinal design in which she repeatedly measured the same group of children over time plotting changes in specific knowledge such as alphabet knowledge, changes in cognitive strategies such as selective attention to highly informative cues, and changes in self-regulation such as self-corrections. She drew on a well-established design to provide the powerful demonstration that these components of development had a pre history (they were emergent before school) and that the developmental properties were best described as a growing orchestration or integration (Clay, 2001).

Her study drew on the early developmental studies such as the Fels Longitudinal study and Gesell’s study at Yale (e.g., Gesell & Ilg, 1946), which asked questions about when changes occur and in what sequences, assuming that the phenomena being measured change in form and function over time. They used the obtained patterns to develop grand theories of development, in Gesell and Ilg’s case a strongly maturational theory which introduced the idea of developmental readiness. Gesell and Ilg produced growth gradients which represented normative trends within which they could identify individual variations. These early studies using the classic designs provided for theorists such as Clay the methodological tools to apply to literacy.

The Development of Literacy? It goes without saying that literacy phenomena are developmental and hence studies which look at the attributes of change over time in form and function are important. Or are they? While it doesn’t need arguing now, this view was not always obvious or accepted. When the idea of emergent literacy developed by Clay (1972) was proposed, it was counter to some developmental views of the time. There was a view that, because children needed to be deliberately taught to read and write at school, examining the attributes of reading and writing over time was not a proper domain for developmentalists. The proper study should be of organically changing aspects of psychological functioning, and hence studies which tried to capture written language development over time were not appropriate (Wolwhill, 1970). Interestingly, this view was despite Gesell and Ilg (1946) having produced growth gradients for “reading behaviour” from 15 months to 5- to 6-years to show that school skills are subject to the principle of developmental readiness, and never the sole product of training or drill.

Clay’s (1972) time-based design helped to establish the developmental nature of literacy. Studies like Clay’s provided an empirical context within which encompassing theories of development could be examined. For example, the longitudinal studies of Ferreiro and Teborosky (1982) interpreted the patterns of changes in writing as evidence for the epigenetic view of developmental change derived from the genetic epistemology of Piaget. The large waves of change in children hypotheses about the nature of the written system could be mapped onto general developmental stages.
Interestingly, the studies by Clay (1972) and further research reviewed by Sulzby and Teale (1991) helped establish an alternative theoretical view to the epigenetic account. Their studies also supported cognitive models of children as thoughtful and strategic. But in addition they provided a means for linking reading and writing to social contexts, notably family or schools. Comparisons between studies illustrated how development could take different forms under different conditions and hence led to theories that predicted multiple pathways of learning. The evidence enabled theorists to establish that socialisation conditions present both at home and at school were central to development, rather than peripheral and provided one foundation for more recent sociocultural theorizing (McNaughton, 1995).

The classic studies over time of this sort were not only longitudinal, there were also cross-sectional designs. The latter are studies in which change over time is captured by taking a slice at several different ages at the same point in time, thus providing a picture of how development changes over these ages, and hence, presumably, over time as children age. Cross sectional studies can also take a form in which only one or two ages are sampled, but these ages are repeatedly sampled over successive years or multiples of years giving a picture of how children of the same age but of different historical cohorts might be similar or different over historical time.

There are variants that are mixtures of the two, not surprisingly termed mixed designs, which provide ways of overcoming readily identified shortcomings of each (Overton & Reese, 1973). For example, longitudinal designs, especially over very long time frames, suffer from problems such as differential drop out and the use of measures that may become outdated. Although cross sectional studies provide a quicker means of plotting shifts over time, they rely on samples of cohorts and because of this do not represent any individual in change over time. They can confute historical change with individual change. Mixed designs capitalize on the strengths of each to provide short-term longitudinal components which enable individual trajectories to be plotted but in a cross sectional format, which means that patterns of change over long time frames can be captured.

The classic designs using longer time frames that have an underlying developmental logic are critical to the study of language arts and continue to be sorely needed. They provide two sorts of important evidence. They yield basic evidence about the nature of development and learning over time. They also provide a means for plotting how wider socialisation changes, including systemic and policy changes at national levels might be associated with changes in children’s literacy levels.

Unraveling the Developmental Threads In terms of the former, the recent analyses by Whitehurst and Lonigan (2001) and by Paris (2005) based on longitudinal studies have enabled important new models of development to be developed and tested. These theorists have provided further evidence to support models of development that are multi componential and assume parallel pathways for the development of components. In the Whitehurst and Lonigan model these are sets of “inside out” (such as sound and print units of knowledge and skills) and “outside in” (knowledge and skills relating to words, semantic units and textual structures) and in the Paris model the equivalent sets are “constrained” and “unconstrained” knowledge and skills.

The assumptions about the developmental relationships and changes are different in these contemporary models and more longitudinal studies are needed to plot conditional relationships and pathways over time using techniques such as structural equation modeling (Whitehurst & Lonigan, 2001) and growth modeling. These can help untangle which assumptions best represent the evidence by explicating the necessary and sufficient relationships between sets of skills and knowledge states. This will enable even better designs for instructional interventions. The need is not just for general theories of literacy but also in those components that have not been adequately mapped. For example, Pearson, Hiebert, and Kamil (2007) have argued that there is a pressing need for studies of growth in vocabulary but that these are dependent on better definitions and measures of vocabulary.

Systemic Changes The second contribution that these designs make is illustrated by cross sectional studies that monitor cohorts of children to examine how a system is performing. National data bases such as the National Assessment of Educational Progress, available since the early 1970s, have enabled long-term monitoring of achievement scores. It is possible to plot patterns of change, in the achievement levels of different groups of students at fourth and eighth grades and relate these to broad policy shifts. For example, Porter & Polikoff (2007) used the data bases to describe how Black students’ performance in reading and writing has been consistently below that of White students, although there have been shifts. In reading there were large gains in Black students’ scores during the 1970s and 1980s. But in the 1990s these gains were reversed and the gaps actually got larger. The picture from 2002 reflects a period of unprecedented federal spending and policy changes through the No Child Left Behind legislation and as yet there is little evidence 3 to 4 years into the programme that gaps have changed markedly (Porter & Polikoff, 2007).

Like the United States, Canada, and other countries, New Zealand also has a national monitoring project which samples a range of reading and writing and oral language tasks for students in year 4 (9-year-olds) and year 9 (14-year-olds) of schooling. Four yearly cycles provide evidence of students’ strengths and weaknesses and how systemic changes such as policy initiatives might impact. From these cycles of assessments it is known that a major literacy strategy commenced in 1998 and focused on school years 1–4 was associated with a positive change in reading accuracy at year 4 (9 years). However, the sampling of tasks
includes comprehension tasks and it also known that there has been little change in comprehension levels for those students most at risk in the system, Māori students (from the indigenous community) and Pasifika students (from recent immigrant or second and older communities from Pacific islands), despite the changes in decoding levels (Lai, McNaughton, Amituanai-Toola, Turner, & Hsiao, 2009).

The Search for Explanations: Unraveling the Variability in Development Over Time

The second set of studies also uses time to study learning and development. Generically labeled “single case” (Hersen & Barlow, 1976), “time series” (Risley & Wolf, 1973), or “single subject” experimental designs (Neuman & McCormick, 1995), these designs use time as a means of identifying causes and hence identifying the means for altering learning and development.

The early foundations for these studies were in the radical behaviourism and operationalism of Skinner (1953) and Sidman (1960), who set out the epistemological bases for the designs. They proposed a disarmingly simple view of cause which equated establishing cause with methodological control. That is, if sources of variability in behaviour could be established so that rates and patterns of learning could be reliably altered under known conditions, then cause had been established.

The original methods, developed with nonhuman subjects, were designed to enable researchers to test and elaborate principles of reinforcement, punishment and stimulus control under tight experimental controls. The basic method involved systematically changing the contingencies of consequences (such as positive reinforcers) and antecedents (such as prompts and instruction) to behaviour, and then plotting the behavioural correlates of these systematic changes against time. If the experimenter could, through systematic replication, reliably and generally repeat the relationships between application and behaviour change, then a robust causal analysis was demonstrated (Sidman, 1960).

The basic methods were elaborated in real world settings using designs that enabled researchers to examine the sources of variability in classrooms and family and other applied settings (Risley & Wolf, 1973). The designs used the idea of baselines established by repeated measurement against which the effects of an introduced intervention could be compared. The core designs included “reversal designs” in which interventions were successively introduced and removed, and “multiple baseline designs” in which interventions were systematically applied in a lagged fashion across multiple subjects or behaviors.

Successful generations of researchers have taken these designs and shown how they can be applied to the language arts (Neuman & McCormick, 1995). They have provided analytic tools to guide the design of effective intervention procedures contributing to a range of productive long-term research programmes such as those of Whitehurst and colleagues (e.g., Whitehurst & Valdez-Menchaca, 1988) who developed the family and preschool language intervention programme of Dialogic Reading, and Brown’s (1997) Reciprocal Teaching and, more broadly, Fostering a Community of Learners programmes.

Optimism and the Solving of Real World Problems

There are two fundamental contributions the second set of studies have made to the range of research endeavors associated with the language arts. One is a commitment to optimism at being able to find and use causes of learning to make a significant difference. A second is the capability to study individuals and groups in ways that enable the identification of sources of variability.

The early applied behaviour analysts were associated with programmes in developmental and educational psychology and had a commitment to change. The optimism was expressed in Baer’s (1973) contribution to a book on methodological issues in life span developmental psychology entitled “The Control of Developmental Process: Why Wait?.” The applied behaviour analysts took literally Hume’s argument in A Treatise of Human nature (1739) that one can’t get an “ought” from an “is.” That is, just because the state of the world is like it is, that is no reason to resign from attempts to change the world into what it ought to be, or perhaps could be.

Was this optimism warranted? As noted above, the portfolio of studies provides an impressive demonstration of their usefulness ranging from precise measurement of instructional effects to contributing to the design of large scale interventions. But there are weaknesses.

The single case designs are essentially quasi-experimental designs because they do not randomise assignment to experimental and control conditions. The single case designs control for issues relating to internal validity by the demonstration and precision of experimental control, but their believability is dependent on how reliably the treatment effect can be shown. In addition, even if internal validity is high there is the issue of external validity, how generalisable are the effects beyond as it were the single case.

Hence, the need to replicate systematically effects and processes is heightened because of the reduced experimental control (Raudenbush, 2005). McCall and Green (2004) argue that in applied developmental contexts, evaluation of programme effects requires a variety of designs including quasi experimental, but our knowledge is dependent on systematic across site analyses. Replication across sites adds to our evaluation of programme effects, particularly when it is inappropriate or premature to conduct experimental randomized designs. Such systematic replication is also needed to determine issues of sustainability (Coburn, 2003) and scaling up (McDonald, Keesler, Kauffman, & Schneider, 2006).

A recent series of large scale interventions that used single case design logic illustrates how the weaknesses in these designs can in part be overcome. The studies were designed to address a long-standing challenge for more
effective teaching in a particular context of culturally and linguistically diverse students (Lai et al., 2009). An intervention model (The Learning Schools Model), which focuses on inquiry processes through which schools can fine tune instruction in reading comprehension, was tested in a systematic replication series across like and unlike clusters of schools (McNaughton & Lai, 2009). The design across three studies uses single case logic within a developmental framework of cross sectional and longitudinal data. The measures of students’ achievement taken at Time 1 generated a cross section of achievement across school year levels (years 4–5–6–7–8), which provides a baseline forecast of what the expected trajectory of development would be if planned interventions had not occurred (Risley & Wolf, 1973). Successive stages of the intervention could then be compared with the baseline forecast and judgments about acceleration that are contextually valid could be made.

This design, which includes replication across cohorts, provides a high degree of both internal and external validity. The internal validity comes from the in-built testing of treatment effects. The external validity comes from the systematic replication across cohorts. A first study was with a cluster of urban primary schools serving culturally and linguistically diverse students from the poorest communities. The systematic replication involved a second cluster of like schools and then a third cluster of unlike (rural and small town) schools. The design has been employed in three studies over 5 years involving up to 10,000 students, over 250 teachers, in 48 schools. The Learning Schools Model has been shown to accelerate, reliably and generally, gains in reading comprehension over 3 years by between 0.5 to 1.0 year in achievement, over and above expected gains during that period.

Despite being quasi-experimental, arguably these designs are well attuned to the nature and vicissitudes of applied settings and are particularly useful in the early stages of developing interventions because of this. Schools are open systems. In school settings researchers have limited control over the many known and unknown influences on teaching and learning, even in the most structured randomised control designs. Over short and unpredictable time frames schools change in features, such as the demographic characteristics of both teachers and students. They change in instructional resources, professional capabilities and in curricula. As was found in the Learning Schools Model studies (Lai et al., 2009), school years are punctuated by summer breaks in which as yet not well-identified family and community practices have effects on achievement. The significance of only some of these potential influences is known, and examining interventions in precise single case studies is a useful tactic.

Of course, randomisation in large-scale experimental studies provides one way of controlling for these potential sources of constraints on replicability. But, even if randomising large numbers of schools is possible in large scale interventions, the logic of a randomisation exercise is flawed as a means for knowing better about the replicability and generality of the interventions in Sidman’s (1960) terms. What we need to know better are the influences on variability in the effectiveness of an intervention, not account for that variability by reducing them to common error variance.

The optimism behind these studies needs to be tempered. There are sources of variability in learning which are very difficult for researchers to change. These are the sources to which Berliner (2006) recently referred when noting the limits to what educational psychological applications can achieve in creating more effective schooling. Public policy, economic and political processes are needed to control variability in such things as poverty, in housing and employment, in access to child care and to well-resourced schools, each of which are sources of differences in achievement between groups in the United States.

**Microgenetic Studies: Strategies and Skills in Dynamic Change**

The third set of studies employs a much briefer time scale than the previous two; typically the scale is of minutes, hours, and days. The idea of studying moment by moment changes in performance has a venerable history which includes Wundt’s (1910) studies in the early 20th century of sensation and perception for which he perfected techniques to study reactions to stimulation over brief time periods.

The newer intensive techniques provide the means for developing and testing theoretical models for children’s complex learning. The central theoretical question asked in these studies is about the processes by which children learn (Siegler, 2000). The basic idea is to observe periods of rapidly changing competence. Children who are on the verge of important developmental changes are provided with exposure to experiences that are predicted to be associated with change, and the course of children’s behaviour is studied intensively.

The models of cognitive skill or expertise assume that moment by moment performance in discrete activities that have known and changing task parameters provide the sites of learning. This is where the dynamic properties of knowledge, strategies and self-regulation can be described and where change and adaptation occur. When engaging in complex learning tasks such as learning to decode narrative texts accurately and fluently, or comprehending an informational text (and thereby potentially acquiring not only new information but also greater skill), new forms of cognitive skill or expertise can emerge.

The method relies on the intensive study of performance in short bursts of real time. By sensitively measuring the properties of children’s performance in relationship to task parameters such as difficulty level and features of the activity setting including sources of external or self-generated feedback, we are in a position to understand the nature of learning and the nature of change in skill. Brown (1997) also used this technique in her studies of communities of learners. Her analyses included studies of the conditions that
promote spontaneous uses of explanations in classrooms and of children’s understanding and use of analogies in reading. In the latter case, 10-year-olds in a laboratory setting were given 4 mini passages to read over 6 days. The passages had analogies to be solved in reference to previous passages. Repeated measurement confirmed that increased problem solution occurred, particularly in the solution of what she termed “deep analogies.”

Microgenetic studies have been significant in developing analyses of intertextuality as a social construction. That is, for understanding how children and teachers jointly construct ideas about connections between texts and how children learn within those constructions (Stuart-Fans & Bloome, 2004). For example, Bloome and Egan-Robertson (1993) analysed a 15-minute segment of instruction in a first-grade classroom. The intensive descriptions of types of linkages made by teachers (such as seeking the correct date to write or giving directions for homework) show how classroom activities are essentially cultural events in the sense of sites where rules and norms for dealing with texts are established and negotiated. This first grade analysis is consistent with Siegler’s (2000) developmental view of these designs in that the transition to school marks a time of rapidly developing competence, the classroom conditions for which can be analysed in considerable detail.

The three sets of studies over time adopt different lenses to study changes in children’s behaviour and to plot the course of development and learning. The lenses provide different contexts for understanding the nature of changes, from macro levels to micro levels. They provide for close description as well as experimentally based explanations. They are important tools in our methods for designing more effective conditions in which children learn.

References


